DOCUMENT RESUME

ED 296 664 HE 021 534

AUTHOR Hewett, Laura; And Others
TITLE Course Evaluations: Are Students' Ratings Dictated by First Impressions?
PUB DATE 88
NOTE 14p.; Paper presented at the Annual Meeting of the Rocky Mountain Psychological Association (Snowbird, VT, April 30, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Faculty; College Students; *Course Evaluation; *Faculty Evaluation; *Grades (Scholastic); Higher Education; *Student Evaluation of Teacher Performance

ABSTRACT

The relationship between students' first impressions of a course (formed on the basis of their initial exam scores) and end-of-term teacher evaluation ratings was studied. University students enrolled in freshman, sophomore, and junior level courses took all tests and completed the instructor evaluation form the week before the fourth quarterly exam. Three 60-question four-alternative multiple choice tests covering the textbook, lecture material, and films were administered before the evaluation in each course. It was found that students tend to evaluate instructors near the end of the course according to their performance on the first examination. The relationship between exam scores and evaluation ratings weakens with successive exams and is weakest with the exam administered just prior to evaluation time. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED296664

Course Evaluations: Are Students' Ratings Dictated by First Impressions?

Laura Hewett, Garvin Chastain, and Steven Thurber

Boise State University

Running head: COURSE EVALUATIONS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

GARVIN CHASTAIN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

HE 021 534

BEST COPY AVAILABLE

Abstract

The purpose of the present study was to investigate the relationship between students' first impressions of a course (formed on the basis of their initial exam scores) and end-of-term teacher evaluation ratings. Results indicate significant correlations between mean Exam 1 scores and mean ratings on all but one evaluation item ($p < .05$ or better), including the overall mean of items. Relationships attenuate on succeeding exams and are weakest for the exam administered and returned to students just prior to evaluations. Evaluation ratings therefore may not be adequate measures of overall teaching effectiveness. These data support the hypothesis that ratings reflect initial impressions of course and instructor as dictated by initial test scores.

Course Evaluations: Are Students' Ratings Dictated by First Impressions?

Many studies have been conducted on the value of students' evaluations of teaching effectiveness. Much of the concern stems from the fact that student evaluations are routinely used in many institutions as part of the basis for faculty promotion and tenure (Soper, 1980).

Are student ratings consistent enough to merit their being used in this way? A great deal of research suggests they are not (Cohen, 1981); student ratings reflect variables such as class size (Gilmore, Swerdlik & Beehr, 1980; Soper, 1980), student's major field of interest (DuCette & Kenney, 1982; Gilmore et al., 1980), instructor popularity, reputation, and ability to entertain the class (Costin, Greenbough & Menges, 1971; Soper, 1980), class level of the course (Aleamoni, 1980), and required versus nonrequired classes (Aleamoni, 1980; Costen et al., 1971).

Brown (1971) suggests students may be rating instructors on how well their teaching style matches the student's perception of the ideal style. Still other investigators suggest evaluation ratings may reflect the purpose students believe the evaluation will serve (Hepworth & Oviatt, 1985).

A literature review indicates the data are not consistent; while the investigators cited found strong relationships, many researchers obtained insignificant correlations. Some of the most inconsistent literature has been written on the Expected Grade Effect. Remmers (1928) reported that the average correlation between grades and ratings was little better than zero -- .070. Rodin and Rodin (1972) took issue with this report, saying "taking the average correlation over traits and instructors is somewhat like characterizing the motion of a swinging pendulum as zero because the two directions cancel each other out". Aleamoni and Thomas (1980), among others, recently found signficant correlations between grades and ratings of the instructor. Marsh (1980) however, suggested that the Expected Grade Effect is "subject to alternative explanations."

Perhaps we should look not at the Expected Grade Effect per se, but at student impressions of the course as dictated by their own test performance. There may be a critical time within the semester when students develop impressions that become evident at evaluation time. The purpose of the present ex-post-facto study was to investigate first impressions, via initial test

performance, as students' possible criterion for evaluating their instructors.

Method

Subjects

University students enrolled in freshman level General Psychology, sophomore level Physiological Psychology and junior level Perception courses with approximate N's of 443, 22 and 11 per course, respectively, served as subjects. Subjects took all tests and were assumed to have completed the instructor evaluation form the week before the fourth quarterly exam.

Materials

Three 60-question four-alternative multiple choice tests over the textbook, lecture material, and films were administered before the evaluation in each course. One test was administered around the end of each quarter of the course.

Evaluation forms required students to rate the instructor on the traits shown in Table 1. Each trait was rated on a seven-point continuum, with one = low and seven = high.

Insert Table 1 about here

Design and Procedures

Six consecutive years of test scores from each of the three courses (1979-1984) were included in the study. All courses were taught by the same instructor.

The freshman and sophomore level courses were required of psychology majors with the former being required by other departments. The sophomore and junior level courses were taken predominately by psychology majors, although the junior level course was not required.

The mean scores for tests one, two and three from each course were compared to the mean evaluation rating from the same course. Mean evaluation ratings for each course were necessarily used instead of individual ratings due to required anonymity of evaluation procedures. However, the mean rating should reflect the impressions of students as a group in each course.

Results

The mean proportion of responses correct on each exam for each class was entered into an analysis of variance. Means for the first, second, and third exams (77.8%, 76.5%, and 76.9%, respectively) did not differ significantly ($F < 1.0$).

Correlation coefficients were computed between mean exam scores and each of the five evaluation items

plus the average of the evaluation ratings for that class, shown as Item 6 (see Table 2). Exam scores from the freshman class were not included in the correlation with Evaluation Item 2 (personal interest in class progress) due to class size (average N = 443).

Insert Table 2 about here

As is shown in Table 2, all but one correlation between mean Exam 1 scores and evaluation items were significant ($p < .05$ or better). Table 2 also shows that correlations between Exam 2 and evaluation items yield fewer (four) significant relationships, and that correlations between Exam 3 and evaluation items yield one significant relationship. Only Evaluation Item 5 (gives reasonable and well-prepared tests) was significantly correlated with all three exams. With the exception of Evaluation Item 2, all correlations attenuate from one exam to the next.

The effect of Exam 1 was partialed out of Exams 2 and 3. Partialing from Exam 2 (see Table 2) reduced three of the five correlations to less than significant values (.36, .41, .16). Those remaining within the significant range were Evaluation Item 5 (gives reasonable and well prepared tests) and the average of

the evaluation ratings. Exam 1 was then partialed out of the Exam 3 - Item 5 correlation (see Table 2) yielding a nonsignificant coefficient (.32).

Discussion

The results of this study support the first-impression hypothesis. Students tend to evaluate instructors near the end of the course according to their performance on the first examination. The relationship between exam scores and evaluation ratings weakens with successive exams and is weakest with the exam administered just prior to evaluation time. The partialing procedures show that the relation between exam performance and ratings is almost fully revealed by considering only scores for Exam 1; that is, considering Exams 2 and 3 adds little to the prediction afforded by considering only Exam 1.

These data may be helpful in explaining the diverse results obtained in Expected Grade Effect studies. Correlations between cumulative test scores and evaluation ratings may hide the first impression effect and yield insignificant relationships in some cases.

Students frequently begin a new course with misconceptions concerning the level of difficulty of the course. The first exam may be a surprise: much

more difficult than the students anticipated. Even if students' grades recover on succeeding exams, the instructor may pay for the students' early misconceptions with low evaluation ratings. Also, if first exam scores are high the instructor may receive high ratings whether or not his or her methods of teaching are effective.

These data suggest that student evaluations do not measure teaching effectiveness over the whole course. Student evaluations may only serve as guidelines for the instructor in preparing his initial course examination.

References

Aleamoni, L. M. (1980). Students can evaluate teaching effectiveness. National Forum, 60, 41.

Aleamoni, L. M., & Thomas, G. S. (1980). Differential relationships of student, instructor, and course characteristics to general and specific items on a course evaluation questionnaire. Teaching of Psychology, 7, 233-235.

Brown, R. (1979). Perceptions of teaching/learning style: The mediating process in student evaluation of instruction. CEDR Quarterly, 12, 16-18.

Cohen, P.A. (1981). Student ratings of instructions and student achievement: A meta-analysis of multisection validity studies. Review of Educational Research, 51, 281-309.

Costin, F., Greenbough, W.T., & Menges, R.J. (1971). Student ratings of college teaching: Reliability, validity and usefulness. Review of Educational Research, 41, 511-535.

Ducette, J., & Kenney, J. (1982). Do grading standards affect student evaluations of teaching? Some new evidence on an old question. Journal of Educational Psychology, 74, 308-314.

Gilmore, D.C., Swerdlik, M.E., & Beehr, T.A. (1980). Effects of class size and college major on student ratings of psychology courses. Teaching of Psychology, 7, 210-214.

Hepworth, D., & Oviatt, B.E. (1985). Using student course evaluations: Findings, issues, and recommendations. Journal of Social Work Education, 3, 105-112.

Marsh, H.W. (1980). The influence of student, course, and instructor characteristics in evaluations of university teaching. American Educational Research Journal, 17, 219-237.

Meredith, G.M. (1982). Grade-related attitude correlates of instructor/course satisfaction among college students. Psychological Reports, 50. 1142.

Remmers, H.H. (1928). The relationship between students' marks and student attitude toward instructors. School and Society, 28, 759-760.

Rodin, M., & Rodin, B. (1972). Student evaluations of teachers. Science, 177, 1164-1166.

Soper, B. (1980). Can students evaluate teaching effectiveness? National Forum, 60, 20,27.

Table 1

Items on the Student Evaluations Being Considered

1. Interests students in subject matter and stimulates thinking
2. Takes an active personal interest in class progress
3. Is well versed in subject and its applications
4. Shows enthusiasm for subject
5. Gives reasonable and well-prepared tests

Table 2

Correlation Coefficients: Mean Exam Scores Versus Mean Ratings on Each Item with Exam 1 Partialed from Significant Correlations for Exams 2 and 3

Exam	Item	N	r	Significance (Two-tailed)	Partialed Coefficients
1	1	18	.5475	.018	-----
1	2[a]	12	.0022	1.000	-----
1	3	18	.6432	.004	-----
1	4	18	.6177	.006	-----
1	5	18	.6735	.002	-----
1	6[b]	18	.6429	.004	-----

(table continues)

Exam	Item	N	r	Significance (Two-tailed)	Partialed Coefficients
2	1	18	.4910	.037	.3655
2	2[a]	12	.3394	.280	-----
2	3	18	.5321	.022	.4065
2	4	18	.3501	.151	.1583
2	5	18	.6411	.004	.5643[c]
2	6[b]	18	.6177	.006	.5446[c]
3	1	18	.2033	.426	-----
3	2[a]	12	.1265	.696	-----
3	3	18	.3205	.192	-----
3	4	18	.1831	.474	-----
3	5	18	.5306	.022	.3190
3	6[b]	18	.3866	.110	-----

[a]Exam scores from the freshman class were not included. [b]Average of evaluation scores. [c]Still significant after partialing.